

Entangled states : Classical versus Quantum

S. Kanmani,
Materials Physics Division, IGCAR,
Kalpakkam, India.
kanmani@igcar.gov.in

July 10, 2009

Abstract

Quantum mechanics of composite systems, gives rise to certain special states called entangled states. A physical system, that is in an entangled state displays an intricate correlation between its subsystems. There are also some composite quantum states (classically correlated states or separable states) that are not entangled. It is generally claimed, often without a rigorous proof to support, that these intricate correlations of an entangled state cannot occur in a classical system. This expository article, provides an elementary proof that entangled states cannot arise in the setting of classical mechanics. In addition, a detailed description of the origin of entanglement in quantum systems is included. The mathematical concepts that are necessary for this purpose are presented. The absence of entanglement in the classical setting is due to the fact that every pure classical state of a composite system is a product state, that is, a tensor product of two pure states of the subsystems. In contrast, there are pure composite quantum states that cannot be expressed in the form of a product state or even by a convex sum of product states. Roughly speaking, this is because classical states are positive valued functions on the phase-space while quantum states are positive linear operators. The structure of the tensor product between two commutative spaces of scalar valued functions is drastically different from that of the tensor product between two non-commutative spaces of linear operators. In other words, entanglement is a non-commutative phenomenon.

Contents

- 1.0 Introduction
- 2.0 States in classical mechanics
- 2.1 Composite classical systems and their states
 - 2.1.1 Product states and separable states
- 3.0 States in quantum mechanics
- 3.1 States as positive operators
- 4.0 Composite quantum systems and their states
- 4.1 Composite quantum systems
 - 4.1.1 Linear functionals and dual vector spaces

- 4.1.2 Bilinear forms
- 4.1.3 Tensor product of vector spaces
- 4.2 States of composite quantum systems
- 4.2.1 Separable states and entangled states.
- 5.0 Appendix (A-E)

In my opinion, the mathematics of last hundred years did not produce anything comparable to quantum theory or general relativity in terms of the resulting change of our total world perception. But I do believe that without the mathematical language physicists could not even say what they were seeing.

- Yuri. I.Manin

1 Introduction

The strategy of decomposing a complex object into simpler parts pervades science. Thus, one tries to understand a quantum mechanical state of a composite system, ¹ comprising of two particles in terms of its constituents, the single particle states. In that context, there arise certain composite states, called entangled states ¹ in which the subsystems display a remarkable correlation between them. For example, knowing the state of one of the particle the state of the other can be predicted with certainty. It is generally said that entanglement is a quantum phenomenon, there by implying such states do not arise in the context of classical mechanics. For example, the article ² states, “ Entanglement is a peculiar property of quantum world that has no classical analog, .. ”. The aim of this article is to provide a pedagogical introduction that clarifies the above statement.

We start with classical mechanics in section 2.0, where the motivation for representing a state as a probability density function on phase-space is given. Section 2.1 considers the cartesian product of phase- spaces as a composite classical system and looks at the nature of product states and separable states. The result that every classical composite state is a separable state and hence is a non-entangled state is obtained in 2.1.1. Section-3 and section-4 are devoted to quantum systems and states. Section 3.0 begins with the notion of a pure state as a vector of unit norm and contains a detailed discussion of mixed states and their mathematical representation. Section 3.1 introduces the notion of density matrices; positive operators with unit trace. Section-4.1 is a self-contained, rigorous introduction to tensor products. Finally, section 4.2 investigates the nature of composite pure states and demonstrates that every quantum mechanical pure state associated with a non-elementary tensor is an entangled state.

Readers interested in quantum information theory and those who wish to go beyond the modest aim of this article may refer to ^{3,4} for more details.

2 States in classical mechanics

In classical mechanics, we represent a state of a particle by specifying a point x_0 in the relevant phase-space X . Recall, a point in a phase-space encodes both position and momentum of the particle. Equivalently, such a state could also be represented by a scalar valued function, $f : X \rightarrow R$, such that $f(x)$ is 1 when $x = x_0 \in X$ and $f(x) = 0$ for all $x \neq x_0$. This function f , can be interpreted as a probability density function defined on the phase-space X . Such a state is called a pure state in the context of classical mechanics or classical statistical mechanics.⁵ A generalisation of this notion, is a probability density function g , defined on the phase-space X , such that $g(x_k) = p_k > 0$, for a finite set of points $\{x_k \in X : 1 \leq k \leq n\}$, such that the sum $\sum_{k=1}^n p_k = 1$. The rest of the points in X , naturally, are assigned the value of zero probability. Such a classical state is called a mixed state.⁵ Thus, in general a classical state is a probability density function defined on a phase-space.

Remark: Mixed states model a situation in which we are not able specify the state sharply by a single point on the phase-space; but can only assure that the system could be in any one of a finite set points, whose probability assignment is non-zero. Observe, that the real system is actually in one of those points. In others words, mixed states model our ignorance of the state of the actual system. This is analogous to the notion of mixed state in quantum mechanics.

Definition-1 A classical state f , associated with a physical system on a phase-space X , is a probability density function on X . That is, a classical state $f : X \rightarrow R$, is a positive valued function such that $f(x) \geq 0$ for every $x \in X$ and $\int_X f dx = 1$.

Note: For the sake of mathematical simplicity, we shall consider only those states f , for which the following set $\{x \in X : f(x) \neq 0\}$, called the support of f , is a finite set. If $\{x_k \in X : 1 \leq k \leq n\}$ is the support of a state f , then the integral $\int_X f dx = 1$, that occurs in the above definition reduces to the sum $\sum_{i=1}^n f(x_i) = 1$.

Definition-2 A classical state f_{x_0} on X , is called a pure state, if the total probability of unity is assigned to a single point $x_0 \in X$. That is, $f_{x_0} : X \rightarrow R$ such that

$$f_{x_0}(x) = \begin{cases} 1 & \text{if } x = x_0 \\ 0 & \text{if } x \neq x_0; \end{cases}$$

In this way every point in the phase-space X gives rise to a pure state.

What is the relation between pure states and mixed states ? We shall show that every mixed state is generated, in a sense, by a set of pure states. First, we observe that the set of all scalar valued functions on the phase-space X , is a vector space. Suppose, f and g are two scalar

valued functions on X . Then one can define their sum $(f + g)$, which is another function on X as follows. Thus $(f + g) : X \rightarrow R$, where $(f + g)(x) := f(x) + g(x)$, for every $x \in X$. Similarly, one can define the multiplication of a scalar $\alpha \in R$ with f , as $(\alpha f) : X \rightarrow R$, where $(\alpha f)(x) := \alpha \times f(x)$, for every $x \in X$. Treating these two operations as vector addition and scalar multiplication respectively, one verifies that the set of all scalar valued functions on X , becomes a vector space. Clearly, classical states are elements of this vector space. Next, we introduce the notion of convex combinations of vectors.

Definition-3 Let $S = \{v_i : 1 \leq i \leq k\}$, be a set of vectors. Then any vector of the form $\sum_{i=1}^k a_i v_i$, where $0 \leq a_i \leq 1$ for $1 \leq i \leq k$ and $\sum_{i=1}^k a_i = 1$ is called a convex combination of vectors from S .

Examples :

1) Let $S = \{v_1, v_2\}$, where v_1, v_2 are two distinct vectors on the plane. Then the set of all convex combinations of v_1 and v_2 is the set $\{pv_1 + (1 - p)v_2 : 0 \leq p \leq 1\}$. Geometrically, this set is the line segment $\overline{v_1 v_2}$, with v_1 and v_2 as their end points.

2) Let $T = \{v_1, v_2, v_3\}$ be a set of three non-collinear vectors on the plane. Then the set of all convex combinations of T , is the set of all the points of the triangular domain, whose vertices are the points v_1, v_2 and v_3 .

Now we are ready for the relation between pure and mixed states.

Proposition-1 Every classical state is either a pure state or a convex combination of pure states. That is, every mixed state is a convex combination of pure states.

Proof: By definition-1 a state f , on a phase-space X is a probability density function on X . By our assumption, the support of f is a finite subset of X . That is, $f(x_i) = p_i > 0$ for a finite subset $\{x_i : 1 \leq i \leq n\}$ of X , and $\sum_{i=1}^n p_i = 1$. Such a function can be expressed as $f = \sum_{i=1}^n p_i f_{x_i}$, where f_{x_i} , represent pure states, for $1 \leq i \leq n$. Recall, the function f_{x_i} , is defined such that $f_{x_i}(x) = 1$, when $x = x_i$ and $f_{x_i}(x) = 0$ for every other $x \in X$. Then, $f(x_k) = \sum_{i=1}^n p_i f_{x_i}(x_k) = \sum_{i=1}^n p_i \delta_{ik} = p_k$, where $1 \leq k \leq n$ and $\delta_{ik} = 1$ if $i = k$ and $\delta_{ik} = 0$ if $i \neq k$. Note, f is a convex combination of pure states. If $n = 1$ then f is a pure state. Thus, by construction any state f , is either a pure state or a convex combination of pure states. A probability density function which assigns a non-zero probability to two or more phase-space points is called a mixed state.

Later, in section-3.1, we shall show that a quantum state is characterised by a positive linear operator with unit trace, called density operator. Observe, the similarities between classical and quantum states. Positive linear operators of quantum mechanics correspond to positive scalar valued functions on phase-space of classical mechanics. Similarly, the condition of unit trace for a quantum state corresponds to the condi-

tion of normalisation; a necessary condition for a positive valued function to be a probability density.

2.1 Composite classical systems and their states:

Consider a particle, called particle-1, whose phase-space is the set X . Similarly, let Y be the phase-space of another particle, called particle-2. The collective system of particle-1 and particle-2, put together constitutes a composite classical system. The phase-space of this composite system is the cartesian product of X with Y , that is, the set $X \times Y = \{(x, y) : x \in X, y \in Y\}$. Clearly, as discussed above, the states of this composite physical system are probability density functions on the set $X \times Y$.

Since every composite state is either a pure state or a convex combination of pure states, we shall look at the pure states first. Any probability density function on $X \times Y$, whose total probability is assigned to a single point, $(x_0, y_0) \in X \times Y$ is a composite pure state. Explicitly, $h_{(x_0, y_0)} : X \times Y \rightarrow R$ is a composite pure state, where

$$[h_{(x_0, y_0)}](x, y) = \begin{cases} 1 & \text{if } (x, y) = (x_0, y_0) \\ 0 & \text{if } (x, y) \neq (x_0, y_0). \end{cases}$$

It is easily verified that this is compatible with proposition-1. In other words, an arbitrary composite mixed state is the same thing as a convex combination of composite pure states of the above form.

2.1.1 Product states and separable states

What is the relation between the pure states of $X \times Y$, the composite system, to the pure states of the subsystems X and Y ? Specifically, let $h_{(x_0, y_0)}$ be a composite pure state as defined above. Then let f_{x_0} and g_{y_0} be the pure states of the subsystems X and Y respectively. Explicitly, $f_{x_0} : X \rightarrow R$, and $g_{y_0} : Y \rightarrow R$, are such that,

$$f_{x_0}(x) = \begin{cases} 1 & \text{if } x = x_0 \\ 0 & \text{if } x \neq x_0. \end{cases}$$

and

$$g_{y_0}(y) = \begin{cases} 1 & \text{if } y = y_0 \\ 0 & \text{if } y \neq y_0. \end{cases}$$

Given two functions $f : X \rightarrow R$, and $g : Y \rightarrow R$ one can define $f \otimes g$, their tensor product as $f \otimes g : X \times Y \rightarrow R$, where $[f \otimes g](x, y) = f(x) \times g(y)$. In the last equality, the product on the right hand side is the product of the real numbers $f(x)$ and $g(y)$. Roughly, this is like multiplying, $P(x)$, a polynomial in the variable x , with $Q(y)$, another polynomial in the

variable y , to get $R(x, y) = P(x) \times Q(y)$, a polynomial in the variables x and y . Essentially, for the space of scalar valued functions, tensor product is the same as the - natural- multiplication of functions as indicated above.

Thus, the tensor product of pure states of the subsystems f_{x_0} and g_{y_0} is of the form $f_{x_0} \otimes g_{y_0} = f_{x_0} \times g_{y_0}$. Clearly, $[f_{x_0} \times g_{y_0}](x, y) = f_{x_0}(x) \times g_{y_0}(y) = \delta_{x_0 x} \times \delta_{y_0 y}$. Hence, this product of two functions takes the value of 1 if and only if $x = x_0$ and $y = y_0$ and takes the value of 0 at all other points. Explicitly,

$f_{x_0} \otimes g_{y_0} : X \times Y \rightarrow R$, such that

$$[f_{x_0} \otimes g_{y_0}](x, y) = f_{x_0}(x) \times g_{y_0}(y) = \begin{cases} 1 & \text{if } (x, y) = (x_0, y_0) \\ 0 & \text{if } (x, y) \neq (x_0, y_0). \end{cases}$$

Note that this is exactly the same as the pure state $h_{(x_0, y_0)}$, of the composite system $X \times Y$. Thus, $f_{x_0} \otimes g_{y_0} = h_{(x_0, y_0)}$. In other words, every pure state of a classical composite system is in the form of a product of pure states of the subsystems. The composite states of the form $f_{x_0} \otimes g_{y_0}$ are called **product states**.

Definition-4 A composite state of the form $f \otimes g$, where f and g are the states of the subsystem is called a **product state**.

Thus we have proved the following proposition.

Proposition-2 Every classical composite pure state is a tensor product of pure states of the subsystems. Thus, every pure state of a composite classical system is a product state.

Note : This is not true for a composite quantum system. In other words, as we shall see, there are pure states in a composite quantum system which cannot be expressed in the form of a product state. In fact, they cannot be even written in the form of a convex combination of product states.

Definition-5 A composite state of the form $\sum_i^n p_i f_i \otimes g_i$, where $\{f_i\}$ and $\{g_i\}$ are the states of the subsystems is called a separable state. Here, $0 \leq p_i \leq 1$ for $1 \leq i \leq n$, and $\sum_{i=1}^n p_i = 1$. If $n = 1$ this becomes a product state. Thus, a separable state is either a product state or a convex combination of product states.

Definition-6 A composite state that is **not** a separable state is called an entangled state. Thus, any state that cannot be expressed as a convex combination of product states is an entangled state.

By proposition-1, every state is either a pure state or a convex combination of pure states. In the case of a composite classical system, every pure state is a product state (cf. Proposition-2). Thus, every classical composite state is either a product state or a convex combination of product

states. Hence, by the definition-5 of separable states, every classical composite state is a separable state. Thus we have the following proposition.

Proposition-3 Every classical composite state is a separable state. Equivalently, there are no entangled states in a classical composite system.

Remark: Given a composite classical state $h = h(x, y)$, on $X \times Y$ one can associate a state $g(y)$, of the subsystem Y in a natural way. This is done by partially integrating the state $h(x, y)$, the probability density, with respect to the variable x , resulting in a marginal probability density $g(y)$ in Y . It is easily verified, that every classical composite pure state thus reduces to a pure state of a subsystem. That is, $\int_X h_{(x_0, y_0)}(x, y) dx = \int_X f_{x_0}(x) \times g_{y_0}(y) dx = g_{y_0}(y)$, where, the states are pure states as defined above. This is not true for a quantum system, where a partial trace¹ of a pure composite state may result in a mixed state of the subsystem. This was first observed by schroedinger. Partial tracing is the quantum analog of partially integrating a composite state over one of the variables of the subsystems.

3 States in quantum mechanics

We shall assume that all our vector spaces are finite dimensional complex vector spaces. Recall, that the quantum mechanical observables associated with position and momentum cannot be modelled⁶ on a finite dimensional vector space. For example, in the context of an electron, only its spin degree of freedom can be modelled on a finite dimensional vector space.

A pure state of a quantum mechanical system is characterised by a vector x of unit norm in a Hilbert space H . As is well known, physical observables are represented by self-adjoint operators acting on that Hilbert space. The expectation value of an observable A , when the system is in a pure state x is given as $\langle x, Ax \rangle$. Here, $\langle u, v \rangle$ denotes the inner product between the vectors u and v of the space H . We shall adopt the convention in which $\langle x, \alpha y \rangle = \alpha \langle x, y \rangle$ and $\langle \alpha x, y \rangle = \bar{\alpha} \langle x, y \rangle$, where $\bar{\alpha}$ denotes the complex conjugate of the complex number α .

Intuitively, a mixed state is a probability density defined on a set of pure states. A simple example of a mixed state is a set containing two pure states $\{x_1, x_2\}$, such that the state x_1 is assigned a probability of p_1 and the state x_2 is assigned the probability $p_2 = 1 - p_1$. Though, the actual system is strictly in only one of those two pure states, we do not know which one of $\{x_1, x_2\}$ is that. Hence, we model this state of uncertainty through a probability distribution on the set of possible pure states. Until we find an appropriate mathematical representation for a general mixed state, we shall denote this mixed state as $S_m = \{(x_1, p_1), (x_2, p_2)\}$; a set of ordered pairs, whose first component is a pure state and the second component is the probability associated with it. The expectation value of an observable A , when the system is in the mixed state S_m , has to be the

weighted sum of $\langle x_1, Ax_1 \rangle$ and $\langle x_2, Ax_2 \rangle$, with their respective probabilities p_1 and p_2 as weights. Thus, the expectation value of an observable A , in the mixed state S_m is $p_1 \langle x_1, Ax_1 \rangle + p_2 \langle x_2, Ax_2 \rangle$, where $p_1 + p_2 = 1$. It is important to understand that a mixed state can not be represented as a vector in H . Suppose we try to represent the mixed state S_m , as a vector $x = p_1 x_1 + p_2 x_2$, where $p_1 + p_2 = 1$; then the expectation value of an observable A , in the state S_m is $\langle x, Ax \rangle = \langle p_1 x_1 + p_2 x_2, A(p_1 x_1 + p_2 x_2) \rangle = p_1^2 \langle x_1, Ax_1 \rangle + p_1 p_2 \langle x_1, Ax_2 \rangle + p_1 p_2 \langle x_2, Ax_1 \rangle + p_2^2 \langle x_2, Ax_2 \rangle = p_1^2 \langle x_1, Ax_1 \rangle + 2p_1 p_2 \text{Re}(\langle x_1, Ax_2 \rangle) + p_2^2 \langle x_2, Ax_2 \rangle$. In the above expression we have made use of the fact that A is self-adjoint and that $\langle u, v \rangle + \langle v, u \rangle$ is equal to two times the real part (denoted as Re) of the complex number $\langle u, v \rangle$. It can be verified that $\langle x, Ax \rangle$ as defined by the expression above is not equal to $p_1 \langle x_1, Ax_1 \rangle + p_2 \langle x_2, Ax_2 \rangle$, the correct expectation value of an observable A in the state S_m . This demonstrates that it is not possible to represent a mixed state as a linear supersposition of pure state vectors.

Hence, our aim is to obtain a mathematical representation of a mixed state that will satisfy the following two conditions. i) Expectation value of an observable A , in the state $S_m = \{(x_1, p_1), (x_2, p_2)\}$, should be $p_1 \langle x_1, Ax_1 \rangle + p_2 \langle x_2, Ax_2 \rangle$. ii) Every mixed state should be a convex combination of pure states.

3.1 States as positive operators

This aim is achieved by representing both pure and mixed states as a particular class of linear operators acting on the Hilbert space H . Suppose S is such an operator representing a quantum state, then the expectation value of an observable A , in the state S is now defined as $\text{Tr}(AS)$, where $\text{Tr}(B)$ denotes the trace of an operator B . In such a generalization, a pure state $x \in H$ is represented as a linear operator $P_x : H \rightarrow H$, defined by its action on $u \in H$ as $P_x(u) = \langle x, u \rangle x$. Then the expectation value of an observable A , in the state P_x is $\text{Tr}(AP_x)$. Now we prove that $\text{Tr}(AP_x) = \langle x, Ax \rangle$ for any pure state x and any observable A as it should be. By definition, trace⁷ of a linear operator T is defined as $\text{Tr}(T) = \sum_{i=1}^n \langle e_i, Te_i \rangle$, where $\{e_i : 1 \leq i \leq n\}$ is any orthonormal basis of H . Given a $x \in H$, it is always possible to find an orthonormal basis $\{e_i : 1 \leq i \leq n\}$, of H in which $e_1 = x$. Then $\text{Tr}(AP_x) = \langle e_1, (AP_x)e_1 \rangle + \sum_{i=2}^n \langle e_i, (AP_x)e_i \rangle = \langle x, (AP_x)x \rangle + \sum_{i=2}^n \langle e_i, (AP_x)e_i \rangle = \langle x, Ax \rangle$. This is because $P_x(x) = \langle x, x \rangle x = \|x\|^2 x = x$ and $P_x(e_i) = 0$, for every $2 \leq i \leq n$, by our choice of orthonormal basis.

By representing pure states x_i as P_{x_i} , the mixed state S_m , could now be expressed as $\rho = p_1 P_{x_1} + p_2 P_{x_2}$, as a convex combination of pure states. Then, the expectation value of an observable A , in the mixed state S_m gets reproduced correctly as $\text{Tr}(A\rho) = \text{Tr}[A(p_1 P_{x_1} + p_2 P_{x_2})] = p_1 \text{Tr}(AP_{x_1}) + p_2 \text{Tr}(AP_{x_2}) = p_1 \langle x_1, Ax_1 \rangle + p_2 \langle x_2, Ax_2 \rangle$. Here, we have used the facts that $\text{Tr}(A+B) = \text{Tr}(A) + \text{Tr}(B)$, $\text{Tr}(\alpha A) = \alpha \text{Tr}(A)$ and the identity $\langle x, Ax \rangle = \text{Tr}(AP_x)$ that we have proved earlier. Thus, we have obtained a mathematical representation of mixed states that is con-

sistent with the two conditions stated above.

Now, we shall show that P_x can be characterised as a self-adjoint, projection operator of rank one. First we shall introduce the notion of rank of a linear operator and show that the pure state P_x is a rank one linear operator. A linear operator is a mapping $T : H \rightarrow H$, such that $T(\alpha u + \beta v) = \alpha T(u) + \beta T(v)$, for every $u, v \in H$ and every $\alpha, \beta \in \mathbb{C}$. The range of a linear operator T , denoted as $\text{range}(T)$ is the set $\{T(x) : x \in H\}$. This set $\text{range}(T)$, for any linear operator T , is a subspace⁷ of H . The rank of a linear operator T , is by definition, the dimension of the range (T). When a linear operator is represented by a matrix, its range is the span of its column (or equivalently row) vectors. Thus, the rank of a matrix M , is the maximal number of linearly independent columns (or equivalently rows) of M .

Recall, that the linear operator $P_x : H \rightarrow H$, that represents a pure state acts on an arbitrary $u \in H$ in the following way. $P_x(u) = \langle x, u \rangle x = zx$, where $\langle x, u \rangle$ denotes the inner product of vector x with u and hence is equal to a complex number z . Thus, P_x maps any vector $u \in H$ into the one dimensional subspace spanned by x . Hence, P_x is a projection operator and the range of P_x is a one dimensional subspace of H . Thus, rank of P_x is one. Since $P_x(u) = \langle x, u \rangle x$ and $P_x(v) = \langle x, v \rangle x$, it follows that P_x is self-adjoint as $\langle v, P_x u \rangle = \langle v, \langle x, u \rangle x \rangle = \langle x, u \rangle \langle v, x \rangle = \langle \langle v, x \rangle x, u \rangle = \langle \langle x, v \rangle x, u \rangle = \langle P_x v, u \rangle$. Similarly, it follows that $P_x P_x = P_x$, because $P_x(P_x(u)) = \langle x, P_x(u) \rangle x = \langle x, \langle x, u \rangle x \rangle x = \langle x, u \rangle \langle x, x \rangle x = \langle x, u \rangle x = P_x(u)$ for every $u \in H$. In Dirac's notation P_x is written as $|x\rangle\langle x|$. We prefer P_x over Dirac's $|x\rangle\langle x|$ as it is convenient in the context of tensor products (cf. appendix-E for more on Dirac's notation). Thus we have a formal definition of quantum states as given below.

Definition 7 A pure state of a quantum mechanical system modelled on a Hilbert space H , is a self-adjoint, rank one projection operator. We shall denote them as P_x , where $x \in H$ and is of unit norm.

Definition 8 A mixed state of a quantum mechanical system modelled on a Hilbert space H is a convex combination pure states. Thus, if ρ is a mixed state then $\rho = \sum_{i=1}^k p_i P_{x_i}$, where $\sum_{i=1}^k p_i = 1$ and P_{x_i} are pure states for $1 \leq i \leq k$.

A classical state is a probability density function and hence is positive valued. We shall show, in a sense, the operators that represent quantum states also have certain positivity property just like the classical states.

Linear operators or equivalently matrices can be thought of as a generalization of complex numbers. Suppose, $T : C \rightarrow C$ is a linear operator acting on the one dimensional complex vector space C . Then, its action on $z \in C$ is as $T(z) = w_T z$, where w_T is a fixed complex number. Equivalently, the 1×1 matrix representation of T is the complex number w_T . Then T^* , the adjoint of T is represented by $\overline{w_T}$, the complex conjugate of w_T . Thus the notion of adjoint is a generalisation of complex conjugation. If T is self-adjoint, then $T = T^*$ or equivalently $\overline{w_T} = w_T$. This

implies that a self-adjoint operator T is represented by a real number w_T . Hence, self-adjoint operators are like real numbers. To summarise, if one thinks of an arbitrary linear operator as a generalized complex number, then self-adjoint operators are like generalised real numbers.

A pure quantum state P_x , being a self-adjoint operator is like a real number. Pushing this analogy between operators and complex numbers further, we claim that P_x is in fact like a positive real number. A complex number z is a positive real number if and only if $z = \overline{w}w$ for some complex number w . Since, adjoint is the appropriate generalisation of complex conjugation, we shall call an operator T to be a positive operator if $T = B^*B$ for some operator B .

Definition 9 An operator $T : H \rightarrow H$ is called a positive operator if $T = B^*B$ for some operator B . Here B^* denotes the adjoint of B .

It is seen immediately that P_x is a positive operator, because $P_x^*P_x = P_xP_x = P_x$. As observed earlier $P_xP_x = P_x$ and $P_x^* = P_x$ as P_x is self-adjoint. Recall an operator T , acting on a Hilbert space is called self-adjoint if $\langle Tu, v \rangle = \langle u, Tv \rangle$ for every $u, v \in H$. In the case of complex vector spaces, there is another definition for self-adjoint operators that is equivalent to this.

Proposition-4 If H is a complex vector space then $T : H \rightarrow H$ is a self-adjoint operator if and only if $\langle Tu, u \rangle = \overline{\langle u, Tu \rangle}$ for every $u \in H$.

Remark: From the property of inner products $\langle Tu, u \rangle$ is the complex conjugate of $\langle u, Tu \rangle$. Thus, in a complex vector space H , an operator T is self-adjoint if and only if $\langle u, Tu \rangle = \overline{\langle u, Tu \rangle}$, or equivalently if and only if $\langle u, Tu \rangle$ is a real number for every $u \in H$.

Proof: (cf. Appendix-A)

Now we record another definition of positive operators, which is equivalent to definition 9 in the context of complex vector spaces.

Definition 10 An operator $T : H \rightarrow H$, on a complex vector space H is positive if $\langle x, Tx \rangle \geq 0$ for every $x \in H$.

Proposition-5 In a complex vector space H , the following two statements about a linear operator $T : H \rightarrow H$ are equivalent.

- 1) $T = B^*B$ for some operator B .
- 2) $\langle x, Tx \rangle \geq 0$ for every $x \in H$.

Proof (cf. Appendix-B)

Proposition-6 A pure state of a quantum mechanical system P_x , is a positive, rank-one operator of unit trace.

Table 1: Analogy between classical and quantum states

Property	Classical	Quantum
State	$f : X \rightarrow R$	$\rho : H \rightarrow H$
Positivity	$f(x) \geq 0 ; x \in X$	$\langle x, \rho x \rangle \geq 0 ; x \in H$
Normalisation	$\int_X f(x) dx = 1$	$\text{Tr } \rho = 1$
Pure state	$\{x \in X f(x) \neq 0\}$ - singleton set	rank of $\rho = 1$

Proof : It has been shown earlier that P_x is a rank-one linear operator. Now, we prove that P_x is a positive operator using definition-10. Since, $\langle u, P_x u \rangle = \langle u, \langle x, u \rangle x \rangle = \langle x, u \rangle \langle u, x \rangle = \langle x, u \rangle \overline{\langle x, u \rangle} \geq 0$, for any $u \in H$, it follows that P_x is a positive operator. Here we have used the properties of inner product and the definition of the linear operator P_x , which acts on $u \in H$ as $P_x(u) = \langle x, u \rangle x$. Now we compute the trace of P_x . By definition, $\text{Tr}(P_x) = \sum_{i=1}^n \langle u_i, P_x u_i \rangle$, where $\{u_i : 1 \leq i \leq n\}$ is any orthonormal basis of H . Choosing, an orthonormal basis of H , in which $u_1 = x$, one gets $\text{Tr}(P_x) = \langle u_1, P_x u_1 \rangle + \langle u_2, P_x u_2 \rangle + \dots + \langle u_n, P_x u_n \rangle = \langle x, P_x x \rangle = \langle x, \langle x, x \rangle x \rangle = \|x\|^4 = 1$ as the later terms vanish and the norm of x being one. Thus, P_x is a positive, rank one operator with unit trace.

Since, a general state is either a pure state or a mixed state, we have the following characterisation of a quantum state.

Proposition-7 A quantum mechanical state is a positive operator of unit trace. Such an operator is called a density operator or matrix.

Proof: A state is either a pure state or a convex combination of pure states. If it is a pure state then by proposition-6 it is a positive operator of unit trace. A mixed state is a convex combination of pure states. Suppose ρ_1 and ρ_2 are two positive operators and $p_1 \rho_1 + p_2 \rho_2$, a convex combination of them. Then, $\langle u, (p_1 \rho_1 + p_2 \rho_2) u \rangle = p_1 \langle u, \rho_1 u \rangle + p_2 \langle u, \rho_2 u \rangle \geq 0$, as ρ_1, ρ_2 are positive operators and p_1, p_2 are positive real numbers. Thus, a convex combination of positive operators, is a positive operator. Hence, a mixed state is a positive operator. Similarly, if $\text{tr}(A) = 1$ and $\text{tr}(B) = 1$ then $\text{tr}(p_1 A + p_2 B) = p_1 \text{tr} A + p_2 \text{tr} B = p_1 + p_2 = 1$. Thus it follows that a convex combination of unit trace operators is an operator of unit trace. Since, pure states are of unit trace it follows that a mixed state, which is a convex combination of pure states is of unit trace as well.

Table-1 displays the analogy between classical states and quantum states.

4 Composite quantum systems and their states

A simple example of composite quantum system is a physical system that consists of two particles. For example, a pair of electrons. The spin degree of freedom of a single electron is modelled on \mathbb{C}^2 , a two dimensional complex vector space. The composite object of two electrons, considering only the spin degree of freedom, is modelled on the vector space $\mathbb{C}^2 \otimes \mathbb{C}^2$, the tensor product space of \mathbb{C}^2 with itself. Hence, one should consider the concept of the tensor product of two vector spaces.

4.1 Composite quantum systems

Now we begin our study of composite quantum systems. Before we define notion of tensor product, we introduce the notions of linear functionals, the dual of a vector space and bilinear functionals or forms.

4.1.1 Linear functionals and dual vector spaces

Given a complex vector space X , consider a complex valued linear mapping ϕ , defined on X . That is, $\phi : X \rightarrow \mathbb{C}$, such that $\phi(\alpha x_1 + \beta x_2) = \alpha \phi(x_1) + \beta \phi(x_2)$ for every $x_1, x_2 \in X$ and $\alpha, \beta \in \mathbb{C}$. Such a ϕ is called a linear form or a linear functional on X . For example, for a fixed $v \in X$, define a linear map $\phi_v : X \rightarrow \mathbb{C}$, where $\phi_v(x) := \langle v, x \rangle$. It can be seen that, $\phi_v(\alpha x_1 + \beta x_2) = \langle v, \alpha x_1 + \beta x_2 \rangle = \alpha \langle v, x_1 \rangle + \beta \langle v, x_2 \rangle = \alpha \phi_v(x_1) + \beta \phi_v(x_2)$, for every $x_1, x_2 \in X$ and $\alpha, \beta \in \mathbb{C}$.

Suppose ϕ and ψ are two linear functionals on X , then their sum $(\phi + \psi)$, is another linear functional on X . This sum is defined as ; $(\phi + \psi)(x) := \phi(x) + \psi(x)$, for every $x \in X$. Similarly, the multiplication of a scalar $\alpha \in \mathbb{C}$ with a linear functional ϕ on X results in a linear functional denoted as $(\alpha\phi)$. This is defined as $(\alpha\phi)(x) := \alpha \times \phi(x)$ for every $x \in X$. With these two operations, as one can verify, the set of all linear functionals on X , becomes a vector space. This is called the dual vector space of X and is denoted as X^* . Note, the zero element of this vector space X^* is a linear functional ϕ_0 , such that $\phi_0(x) = 0 \in \mathbb{C}$ for every $x \in X$. Often, we shall denote the zero linear functional by 0. If ϕ is a non-zero linear functional, then there is a $x \in X$ such that $\phi(x) \neq 0$. In particular, if $\phi(x) = 0$ for every linear functional $\phi \in X^*$, then $x = 0$. Let $E = \{e_1, e_2, \dots, e_n\}$ be a basis of X . Then a linear map T on X gets completely specified by the values $\{T(e_k) : e_k \in E\}$. For example, if $\{e_1, e_2\}$ is a basis of a two dimensional vector space X , then there is a unique linear functional $\phi \in X^*$, such that $\phi(e_1) = 1$ and $\phi(e_2) = 0$. Later, we shall make use of such linear functionals.

4.1.2 Bilinear forms

We shall define a tensor product space as a dual vector space of the space of bilinear forms. Hence, we shall introduce the notion of a bilinear form. Suppose X and Y are two vector spaces. Then a complex valued function f , defined on $X \times Y$ is called a bilinear form if it satisfies the following conditions. 1) $f(\alpha x_1 + \beta x_2, y) = \alpha f(x_1, y) + \beta f(x_2, y)$ for every $x_1, x_2 \in X$, $y \in Y$ and $\alpha, \beta \in \mathbb{C}$ and 2) $f(x, \alpha y_1 + \beta y_2) = \alpha f(x, y_1) + \beta f(x, y_2)$ for every $y_1, y_2 \in Y$, $x \in X$ and $\alpha, \beta \in \mathbb{C}$. That is, f is a function of two (vector) variables such that f acts as a linear map in each variable when the other variable is fixed. Now we shall look at an example of a bilinear form. Let X be a vector space and X^* its dual. Then the map $b : X \times X^* \rightarrow \mathbb{C}$, where $b(x, \phi) := \phi(x)$, $x \in X$, $\phi \in X^*$, is a bilinear form.

Note, that a bilinear form is not a linear map. Clearly, the domain of a bilinear form, that is, the set $X \times Y$ is not even a vector space. However, the set of all bilinear forms from $X \times Y$ to \mathbb{C} , is a vector space. The sum of two bilinear forms and the multiplication of a complex scalar with a bilinear form are defined pointwise, just as we did in the case of linear functionals. For example, if f and g are two bilinear forms on $X \times Y$, then $(f + g)(x, y) := f(x, y) + g(x, y)$, for every $(x, y) \in X \times Y$. Similarly, one can define the multiplication of a complex scalar, with a bilinear form. We shall denote this vector space, that is, the vector space of all bilinear forms on $X \times Y$ as $B(X \times Y)$.

4.1.3 Tensor product of vector spaces

The notion of tensor product involves many abstract concepts. First of all, keep in mind that the symbol $X \otimes Y$, stands for a vector space. The symbols X and Y in $X \otimes Y$, remind us that it has been created, crudely speaking, by a sort of product or multiplication of two vector spaces X and Y . The elements of $X \otimes Y$ are vectors. However, to emphasize the fact that these elements were obtained by the special process of - tensor product - of two vector spaces, we shall call them tensors. The space $X \otimes Y$, contains some elements that can be considered as if they were obtained by multiplying an element $x \in X$ with another element $y \in Y$. We shall denote such an element as $x \otimes y$. Such elements are called elementary tensors. Infact, every element in $X \otimes Y$ is a sum of elementary tensors. Note, in the context of the symbol $x \otimes y$, that $x \in X$, $y \in Y$ and $x \otimes y \in X \otimes Y$.

Formally, the tensor product, $X \otimes Y$, of the vector spaces X and Y is defined as the dual space of the vector space of bilinear forms $B(X \times Y)$. That is, if $\tau \in X \otimes Y$, then τ is a linear functional from the vector space of $B(X \times Y)$ to the space of complex numbers. Specifically, $\tau : B(X \times Y) \rightarrow \mathbb{C}$, is defined such that $\tau(\alpha f_1 + \beta f_2) = \alpha \tau(f_1) + \beta \tau(f_2)$, for every bilinear form $f_1, f_2 \in B(X \times Y)$ and $\alpha, \beta \in \mathbb{C}$.

Hence, if $x \in X$ and $y \in Y$, then the symbol $x \otimes y$, as we defined above, denotes a linear functional on $B(X \times Y)$. That is, $x \otimes y$ stands for a linear

map from the vector space of $B(X \times Y)$ to the space of complex numbers. Formally, $x \otimes y : B(X \times Y) \rightarrow \mathbb{C}$ and the action of $x \otimes y$ on a bilinear form $f \in B(X \times Y)$ is defined as ; $(x \otimes y)(f) := f(x, y)$. If $x' \otimes y'$ is another linear functional acting on $B(X, Y)$, then their sum denoted as $x' \otimes y' + x \otimes y$ is defined as follows; $(x' \otimes y' + x \otimes y)f = (x' \otimes y')f + (x \otimes y)f = f(x', y') + f(x, y)$ for every bilinear form $f \in B(X \times Y)$. Similarly, the multiplication of a complex scalar α with a linear functional results in another linear functional. This is done by defining it as $(\alpha(x \otimes y))(f) := \alpha \times (x \otimes y)(f) = \alpha \times f(x, y)$ for every $\alpha \in \mathbb{C}$. Thus, $X \otimes Y$, is the vector space of all linear functionals spanned by the functionals of the form $x \otimes y$. Tensors of the form $x \otimes y$, are called elementary tensors. Formally, $X \otimes Y = \text{span} \{x \otimes y : x \in X, y \in Y\}$.

Definition-11 The tensors of the form $x \otimes y$, where $x \in X$ and $y \in Y$ are called elementary tensors. They span the entire tensor product space $X \otimes Y$.

Caution : The set of all elementary tensors is not a linearly independent set; for the reason that there are too many of them. Hence, even though they span the entire vector space $X \otimes Y$, they do not constitute a basis. One important consequence of this that representation of an arbitrary tensor in terms of elementary tensors is not unique. Two different looking tensors may actually turn out to be equal !

The elementary tensor of the form $(x + x') \otimes y$ acts on a bilinear form f in the following way. $[(x + x') \otimes y](f) = f(x + x', y) = f(x, y) + f(x', y) = (x \otimes y)(f) + (x' \otimes y)(f) = [(x \otimes y) + (x' \otimes y)](f)$. Since, this is valid for every bilinear form f , it follows that, $(x + x') \otimes y = x \otimes y + x' \otimes y$. Similar reasoning leads to the following list of identities.

1. $(x_1 + x_2) \otimes y = x_1 \otimes y + x_2 \otimes y$
2. $x \otimes (y_1 + y_2) = x \otimes y_1 + x \otimes y_2$
3. $(\alpha x) \otimes y = \alpha(x \otimes y) = x \otimes (\alpha y)$
4. $0_X \otimes y = x \otimes 0_Y = 0_{X \otimes Y}$

where $x, x_1, x_2 \in X$; $y, y_1, y_2 \in Y$ and α is a complex number. The symbols, $0_X, 0_Y$ and $0_{X \otimes Y}$ denote the null vectors of the vector spaces X, Y and $X \otimes Y$ respectively.

These properties are summarised by saying that the tensor product \otimes , is a bilinear map from $X \times Y$ to $X \otimes Y$. Note, this map takes the pair (x, y) to $x \otimes y$. From an abstract⁹ point of view this is the most important bilinear map for the pair of vector spaces (X, Y) . If you call this bilinear map b , then $b : X \times Y \rightarrow X \otimes Y$, and $b(x, y) = x \otimes y$. Now given any vector space W and a bilinear map $f : X \times Y \rightarrow W$, there is a unique linear map $T_f : X \otimes Y \rightarrow W$ such that f can be factored as, $f = T_f \circ b$. That is, $f(x, y) = T_f \circ b(x, y) = T_f(x \otimes y)$, $x \in X$, $y \in Y$. Essentially, the pair $(b, X \otimes Y)$, -converts- bilinear maps on $X \times Y$, into linear maps on $X \otimes Y$.

Suppose, x_1 and x_2 are two linearly dependent vectors in X and y_1 , and y_2 are arbitrary vectors in Y , then the tensor of the form $t = x_1 \otimes y_1 + x_2 \otimes y_2$ is actually an elementary tensor. This is because, $x_1 \otimes y_1 + x_2 \otimes y_2 = (\alpha x) \otimes y_1 + (\beta x) \otimes y_2 = x \otimes (\alpha y_1) + x \otimes (\beta y_2) = x \otimes (\alpha y_1 + \beta y_2) = x \otimes y$ where $y = \alpha y_1 + \beta y_2$. Here, we have made use of the fact that $\{x_1, x_2\}$ is a linearly dependent set and hence $x_1 = \alpha x$ and $x_2 = \beta x$ for some $x \in H_1$. The rest of the steps follow from the bilinear properties of the tensor product listed above. Thus, a linear combination of elementary tensors is a non-elementary tensor if and only if it cannot be reduced to an elementary tensor as we have just demonstrated. An example of a non-elementary tensor is $\tau = u_1 \otimes v_1 + u_2 \otimes v_2$, where $\{u_1, u_2\}$ is a linearly independent set in H_1 and $\{v_1, v_2\}$ is a linearly independent set in H_2 . Such a τ can never be written in the form of $x \otimes y$. This important fact is also crucial for our final result.

Proposition - 8 Let $X \otimes Y$ be the tensor product of vector spaces X and Y . Suppose, $\{u_1, u_2\}$ is a linearly independent subset of X and $\{v_1, v_2\}$ is a linearly independent subset of Y . Then a tensor of the form $u_1 \otimes v_1 + u_2 \otimes v_2$ is not equal to $u \otimes v$ for any $u \in X$ and $v \in Y$. Hence, a non-elementary tensor can never be expressed as a scalar multiple of an elementary tensor.

Proof : (cf. Appendix-C)

With these tools we begin our study of composite quantum states.

4.2 States of composite quantum systems

Consider a composite quantum mechanical system that consist of two particles, say, particle-1 and particle-2. If the particle-1, as an individual entity, was modelled on a Hilbert space H_1 and the particle-2, as an individual entity, was modelled on a Hilbert space H_2 , then the composite system is modelled on the tensor product space of $H_1 \otimes H_2$. Hence, the composite states (both pure and mixed) are operators that act on $H_1 \otimes H_2$. First we shall look at pure states. Clearly, by definition-7, a pure state of this composite system is a self-adjoint, rank-one projection operator acting on $H_1 \otimes H_2$. As before, we shall denote it by P_t , where $t \in H_1 \otimes H_2$. Observe, that t could either be an elementary tensor or a non-elementary tensor. First we look at the case of elementary tensor.

Proposition - 9 A pure state $P_t : H_1 \otimes H_2 \rightarrow H_1 \otimes H_2$, of a composite quantum system on $H_1 \otimes H_2$, where $t = x \otimes y$, an elementary tensor is a tensor product of pure states of the subsystems. Equivalently, $P_{x \otimes y} = P_x \otimes P_y$. Here, P_x and P_y are the pure states of the subsystems on H_1 and on H_2 respectively.

Note As such a state, is in the form of a (tensor) product of states of subsystems, it is called a product state. Observe, this is a tensor product of operators. The set of all linear operators or matrices on a vector space

itself is a vector space. Hence, tensor product of two such spaces of operators is well defined. For example, if M_2 denotes the vector space of 2×2 complex matrices, then $M_2 \otimes M_2$ denotes the tensor product of M_2 with itself.

Proof:

Let $t = x \otimes y$, be an elementary tensor in $H_1 \otimes H_2$. Then $P_t : H_1 \otimes H_2 \rightarrow H_1 \otimes H_2$ acts on $\tau \in H_1 \otimes H_2$ in the following way. $P_t(\tau) = \langle \langle t, \tau \rangle \rangle t$. Here $\langle \langle \tau, t \rangle \rangle$ denotes the inner product of the tensor product space $H_1 \otimes H_2$. This innerproduct is defined as

$$\langle \langle u_1 \otimes u_2, v_1 \otimes v_2 \rangle \rangle = \langle u_1, v_1 \rangle_{H_1} \times \langle u_2, v_2 \rangle_{H_2}$$

for elementary tensors and is extended to arbitrary tensors using the well known properties of inner product. In the following we shall suppress the subscripts H_i , on the inner products $\langle \cdot, \cdot \rangle_{H_i}$ for the sake of readability. Let $\tau = u \otimes v$, then

$$\begin{aligned} P_t(\tau) &= P_{x \otimes y}(u \otimes v) \\ &= \langle \langle x \otimes y, u \otimes v \rangle \rangle [x \otimes y] \text{ (definition of } P_{x \otimes y} \text{)} \\ &= \langle x, u \rangle \langle y, v \rangle [x \otimes y] \text{ (definition of } \langle \langle \cdot, \cdot \rangle \rangle \text{)} \\ &= [\langle x, u \rangle x \otimes \langle y, v \rangle y] \text{ (using the bilinearity of } \otimes \text{)} \\ &= P_x u \otimes P_y v = [P_x \otimes P_y](u \otimes v) \end{aligned}$$

Thus, $P_{x \otimes y}(u \otimes v) = [P_x \otimes P_y](u \otimes v)$ for an arbitrary elementary tensor $(u \otimes v)$. Since, $P_{x \otimes y}$ is a linear operator, this equality extends to non-elementary tensors as well. Thus, $P_{x \otimes y}(\tau) = P_x \otimes P_y(\tau)$, for an arbitrary tensor $\tau \in H_1 \otimes H_2$. Hence, $P_{x \otimes y} = P_x \otimes P_y$. This proves proposition-9.

4.2.1 Separable states and entangled states

When a composite system is in a product state $P_x \otimes P_y$, one says that particle-1 is in the state P_x of the subsystem H_1 and particle-2 is in the state P_y of the subsystem H_2 . This implies, that these two particles act independent of each other. That is, there is no correlation between them. This situation is analogous to the case in probability theory, where two random variables x and y are said to be independent if their composite probability density $\phi(x, y)$, can be written as a product of individual densities, say, as $\phi(x, y) = \phi_1(x) \times \phi_2(y)$. In fact, not only a product state but any convex combination of such product states also do not have a strong correlation between the subsystems. Such a state is called a separable state.

Definition-12 A state ρ of a composite system $H_1 \otimes H_2$ is said to be a separable state if it can be expressed as $\sum_{i=1}^m p_i (P_{x_i} \otimes P_{y_i}) = \sum_{i=1}^m p_i P_{x_i \otimes y_i}$, where P_{x_i} and P_{y_i} are the pure states of the subsystems H_1 and H_2

respectively for $1 \leq i \leq m$. Here, $p_i \geq 0$ for $1 \leq i \leq m$ and $\sum_{i=1}^m p_i = 1$. Note, $P_{x_i} \otimes P_{y_i} = P_{x_i \otimes y_i}$ represents a pure state of the composite system associated with the elementary tensor $x_i \otimes y_i$. Observe, when $m = 1$, a separable state becomes a product state. Thus, a separable state is a convex combination of product states, that is, states of the form $P_{x_i} \otimes P_{y_i}$.

Proposition-10 A pure state P_t , of the composite system $H_1 \otimes H_2$, where t is an elementary tensor, is a separable state.

Proof : This is because by Proposition-9, every pure state P_t , where t as an elementary tensor is equal to $P_x \otimes P_y$, for some $x \in H_1$, $y \in H_2$. Hence, such a state is a separable state.

Separable states are also called as classically correlated states¹⁰. This is justified because, as we saw in section 2.1, every classical composite state is in the form of a separable state. Right now it is not at all obvious that there are states that are not separable. One expects non-separable states to have certain degree of correlation between its subsystems. A composite states that is not in the form of a separable state is called an entangled state.

Definition-13 A composite state that is not separable is called an entangled state.

Before we get to look at entangled states, we need one more result on separable states. This result is known as the range criterion in quantum information theory.

Proposition -11 The range of a separable state $\rho_s : H_1 \otimes H_2 \rightarrow H_1 \otimes H_2$, which is a subspace of $H_1 \otimes H_2$, is spanned by elementary tensors. That is, the subspace $\text{range} \rho_s$, has a basis that consists entirely of elementary tensors. For example, if $\rho_s = p_1 P_{x_1 \otimes y_1} + p_2 P_{x_2 \otimes y_2}$, then $\text{range}(\rho_s) = \text{span}\{x_1 \otimes y_1, x_2 \otimes y_2\}$.

Proof : (cf. Appendix-D)

There are plenty of pure states in a composite system, which are of the form P_τ , where τ is a non-elementary tensor. This is the case, for example, if $\tau = u_1 \otimes v_1 + u_2 \otimes v_2$, where $\{u_1, u_2\}$ is a linearly independent set in H_1 and $\{v_1, v_2\}$ is a linearly independent set in H_2 .

Now we prove that every pure state that is associated with a non-elementary tensor is not a separable state.

Proposition-12 A composite pure state P_τ , a rank one, self-adjoint, projection operator acting on $H_1 \otimes H_2$, where τ is a non-elementary tensor in $H_1 \otimes H_2$ represents an entangled state.

Proof: Assume the contrary. That is, let $P_\tau = \rho_s$, where ρ_s is a separable state. Let $\tau = u_1 \otimes v_1 + u_2 \otimes v_2$, be the non-elementary tensor. Then $\{u_1, u_2\}$ is a linearly independent set in H_1 and $\{v_1, v_2\}$ is a linearly independent set in H_2 . It follows, that the range of P_τ is equal to the range of

ρ_s . P_τ , being a pure state has a one dimensional range spanned by τ . That is, the range of P_τ is the set $\{\alpha\tau = \alpha(u_1 \otimes v_1 + u_2 \otimes v_2) : \alpha \in \mathbb{C}\}$, a one dimensional subspace of $H_1 \otimes H_2$. By proposition- 11, the range of a separable state ρ_s is spanned by elementary tensors. Since, $P_\tau = \rho_s$, the range of ρ_s is also a one dimensional subspace spanned by an elementary tensor, say, $x \otimes y$. Thus, the $\text{Range}(P_\tau) = \{\alpha\tau = \alpha(u_1 \otimes v_1 + u_2 \otimes v_2) : \alpha \in \mathbb{C}\} = \text{Range}(\rho_s) = \text{span}\{x \otimes y\} = \{\beta(x \otimes y) : \beta \in \mathbb{C}\}$. By Proposition- 8, it is not possible to express a non-elementary tensor as a scalar multiple of an elementary tensor. Thus, we have reached a contradiction. Hence, $P_\tau \neq \rho_s$, for any separable ρ_s . So we conclude that P_τ , when τ is a non-elementary tensor is an entangled state.

In the case of classical states, every composite pure state turned out to be a product of pure states of subsystems, called a product state and hence a non-entangled state. Moreover, as every state is a convex combination of pure states, all states turn out to be convex combination of such product states, that is, non-entangled states. However, as we have realised, the pure states of composite quantum systems that are associated with non-elementary tensors are entangled. In fact, there are also mixed states which are entangled in the case of quantum mechanics. In contrast, no classical state, either pure or mixed is an entangled state.

Acknowledgement The author thanks the members of the theory group at MSD, IGCAR for their active participation in many discussions on entanglement and the organisers of the conference, " Entanglement in quantum condensed matter," held at the Institute of Mathematical sciences, Chennai, during 17-29, November, 2008.

5 Appendix

Appendix-A

Proposition-4 If H is a complex vector space then $T : H \rightarrow H$ is a self-adjoint operator if and only if $\langle Tu, u \rangle = \langle u, Tu \rangle$ for every $u \in H$.

Proof: Let H be a complex vector space. We have to show that $\langle Tu, u \rangle = \langle u, Tu \rangle$ for every $u \in H$ is equivalent to $\langle Tu, v \rangle = \langle u, Tv \rangle$ for every $u, v \in H$. Suppose, $\langle Tu, v \rangle = \langle u, Tv \rangle$, for every $u, v \in H$ then it is obvious by putting $u = v$ that $\langle Tu, u \rangle = \langle u, Tu \rangle = \overline{\langle Tu, u \rangle}$ for every $u \in H$. In the other direction, suppose $\langle Tx, x \rangle = \langle x, Tx \rangle$ for every $x \in H$ then $\langle T(u + \alpha v), (u + \alpha v) \rangle = \langle (u + \alpha v), T(u + \alpha v) \rangle$ for every $u, v \in H$ and $\alpha \in \mathbb{C}$. Expanding the above expression leads to the equality $\langle u, T\alpha v \rangle + \langle \alpha v, Tu \rangle = \langle Tu, \alpha v \rangle + \langle T\alpha v, u \rangle$. Which implies $\text{Im}(\alpha \langle u, Tv \rangle) = \text{Im}(\alpha \langle Tu, v \rangle)$. We use $\text{Im}(z)$ and $\text{Re}(z)$ to denote the imaginary and real part of complex number z respectively. The equality being valid for every complex number α ; Choosing $\alpha = i$, where $i^2 = -1$, it follows $\text{Re}(\langle u, Tv \rangle) = \text{Re}(\langle Tu, v \rangle)$ and choosing $\alpha = 1$, it follows $\text{Im}(\langle u, Tv \rangle) = \text{Im}(\langle Tu, v \rangle)$. Thus $\langle u, Tv \rangle = \langle Tu, v \rangle$.

Remark: Proposition-4 cannot be extended to real vector spaces. For example, the 2×2 real matrix A , with $A_{1,1} = A_{2,2} = 1, A_{1,2} = 2$ and $A_{2,1} = 0$, considered as an operator acting on R^2 is not self-adjoint, even though $\langle x, Ax \rangle$ is a real number for every $x \in R^2$.

Appendix-B

Proposition-5 In a complex vector space H , the following two statements about a linear operator $T : H \rightarrow H$ are equivalent.

- 1) $T = B^*B$ for some operator B .
- 2) $\langle x, Tx \rangle \geq 0$ for every $x \in H$.

If $T = B^*B$, then $\langle x, Tx \rangle = \langle x, B^*Bx \rangle = \langle Bx, Bx \rangle = \|Bx\|^2 \geq 0$, by the axioms of norm. In the other direction, if $\langle x, Tx \rangle \geq 0$ for every $x \in H$ then by proposition-4 it follows that T is self-adjoint. We claim that all the eigen values of T are non-negative. Suppose u is an eigenvector of T , with eigenvalue λ , then $\langle u, Tu \rangle = \langle u, \lambda u \rangle = \lambda \langle u, u \rangle \geq 0$, which implies λ and hence all the eigenvalues of T are positive. Since T being self-adjoint the eigenvectors of T form a basis of H . Then, such a T can be expressed, in the basis consisting of its eigenvectors, as a diagonal matrix with its non-negative eigenvalues λ_i as diagonal elements. By a diagonal matrix we mean a matrix whose non-diagonal entries are all zero. We denote the matrix that represents the operator T as $[T]$. Thus we have $[T] = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ where $\lambda_i \geq 0$ for $1 \leq i \leq n$. Now one can write $[T] = [B^*][B]$, where $[B] = \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n})$. This completes the proof.

Appendix-C

Proposition-8 Let $X \otimes Y$ be the tensor product of vector spaces X and Y . Suppose, $\{u_1, u_2\}$ is a linearly independent subset of X and $\{v_1, v_2\}$ is a linearly independent subset of Y . Then a tensor of the form $u_1 \otimes v_1 + u_2 \otimes v_2$ is not equal to $u \otimes v$ for any $u \in X$ and $v \in Y$.

Proof : We have to show that $u_1 \otimes v_1 + u_2 \otimes v_2 \neq u \otimes v$ for any $u \otimes v \in X \otimes Y$. Equivalently, $u_1 \otimes v_1 + u_2 \otimes v_2 - u \otimes v \neq 0$ for any $u \otimes v \in X \otimes Y$. We assume the contrary and reach a contradiction. Let $u_1 \otimes v_1 + u_2 \otimes v_2 - u \otimes v = 0$. Recall, if a tensor $\tau \in X \otimes Y$ is zero then $\tau(f) = 0$, for every bilinear form $f : X \times Y \rightarrow \mathbb{C}$. Specifically, $\tau(f)$ is defined such that, if $\tau = u \otimes v$, then $\tau(f) = (u \otimes v)(f) = f(u, v)$. If $\tau = u_1 \otimes v_1 + u_2 \otimes v_2$, then $\tau(f) = (u_1 \otimes v_1 + u_2 \otimes v_2)(f) = (u_1 \otimes v_1)(f) + (u_2 \otimes v_2)(f) = f(u_1, v_1) + f(u_2, v_2)$, for every bilinear form f . So if $(u_1 \otimes v_1 + u_2 \otimes v_2 - u \otimes v)$ is a zero tensor then $(u_1 \otimes v_1 + u_2 \otimes v_2 - u \otimes v)(f) = 0$ for every bilinear form f . Note, $(u_1 \otimes v_1 + u_2 \otimes v_2 - u \otimes v)(f) = (u_1 \otimes v_1)(f) + (u_2 \otimes v_2)(f) - (u \otimes v)(f) = f(u_1 \otimes v_1) + f(u_2 \otimes v_2) - f(u \otimes v) = f(u_1, v_1) + f(u_2, v_2) - f(u, v)$, for every bilinear form f .

Now we construct some bilinear forms, using the linear functionals that act on X and Y . We shall use the symbol ϕ and ψ for an arbitrary linear

functional in X^* and Y^* respectively. Observe, if $\phi \in X^*$ and $\psi \in Y^*$, then $\phi : X \rightarrow \mathbb{C}$ and $\psi : Y \rightarrow \mathbb{C}$. Then we define a bilinear form $\phi \times \psi$, on $X \times Y$ such that $(\phi \times \psi)(x, y) = \phi(x) \times \psi(y)$, $x \in X$, $y \in Y$. Since, $\{u_1, u_2\}$ is a linearly independent set in X , we can construct an ordered basis of X , which includes u_1 and u_2 as its first two elements. That is, $\{u_1, u_2, u_3, \dots, u_n\}$ is a basis of X . Then, let $\phi_1 : X \rightarrow \mathbb{C}$ be a linear functional in the dual space X^* , such that $\phi_1(u_1) = 1$ and $\phi_1(u_k) = 0$ for all $k \neq 1$. Such a linear functional always exist as we discussed above in section 4.1 on dual spaces. Similarly, as $\{v_1, v_2\}$ is a linearly independent set in Y , one can construct an ordered basis of Y , which includes v_1 and v_2 as its first two elements. That is, $\{v_1, v_2, v_3, \dots, v_m\}$ is a basis of Y . Then, let $\psi_2 : Y \rightarrow \mathbb{C}$ be an element in the dual space Y^* , such that $\psi_2(v_2) = 1$ and $\psi_2(v_k) = 0$ for all $k \neq 2$.

Step-1 We claim that v_1 and v are linearly dependent. Consider, a bilinear form f , such that $f(x, y) = \phi_1(x) \times \psi(y)$, $x \in X$, $y \in Y$, where ϕ_1 is the particular linear functional as defined above and ψ is an arbitrary linear functional in Y^* . We shall denote this bilinear form as $\phi_1 \times \psi$. Since, $(u_1 \otimes v_1 + u_2 \otimes v_2 - u \otimes v)(f) = 0$, for every bilinear form f , it follows $(u_1 \otimes v_1 + u_2 \otimes v_2 - u \otimes v)(\phi_1 \times \psi) = 0$. Which implies $(u_1 \otimes v_1)(\phi_1 \times \psi) + (u_2 \otimes v_2)(\phi_1 \times \psi) - (u \otimes v)(\phi_1 \times \psi) = \phi_1(u_1)\psi(v_1) + \phi_1(u_2)\psi(v_2) - \phi_1(u)\psi(v) = 0$. By the definition of ϕ_1 , $\phi_1(u_1) = 1$ and $\phi_1(u_2) = 0$. Hence, we have $\psi(v_1) - \phi_1(u)\psi(v) = \psi(v_1 - \phi_1(u)v) = 0$. Since, ψ is an arbitrary linear functional in Y^* , it follows $v_1 - \phi_1(u)v = 0$. Here, we are using the fact (cf. Section 4.1) that if $\psi(y) = 0$ for every $\psi \in Y^*$ then $y = 0$. Since, $v_1 - \phi_1(u)v = 0$, we conclude that v_1 and v are linearly dependent. Hence, $\psi_2(v) = \psi_2(\alpha v_1) = \alpha \psi_2(v_1) = 0$, as by definition $\psi_2(v_1) = 0$.

Step-2 Now we claim u_2 is zero, which is in contradiction to the fact that $\{u_1, u_2\}$ is linearly independent. Recall, any set that contains a null vector is linearly dependent. Consider a bilinear form f , such that $f = \phi \times \psi_2$, where ϕ is an arbitrary linear functional in X^* and ψ_2 is the specific linear functional in Y^* , that was defined above. Then, $f(x, y) = \phi(x)\psi_2(y)$, $x \in X$, $y \in Y$. Since, $(u_1 \otimes v_1 + u_2 \otimes v_2 - u \otimes v)(f) = 0$, for every bilinear form f , it follows $(u_1 \otimes v_1 + u_2 \otimes v_2 - u \otimes v)(\phi \times \psi_2) = (u_1 \otimes v_1)(\phi \times \psi_2) + (u_2 \otimes v_2)(\phi \times \psi_2) - (u \otimes v)(\phi \times \psi_2) = 0$. Which implies, $\phi(u_1)\psi_2(v_1) + \phi(u_2)\psi_2(v_2) - \phi(u)\psi_2(v) = \phi(u_2) = 0$. Here we have used the properties of ψ_2 that $\psi_2(v_2) = 1$, $\psi_2(v_1) = 0$, and the fact $\psi_2(v) = 0$, which was obtained at the end of step-1. Since, $\phi(u_2) = 0$ and ϕ is an arbitrary linear functional it follows that $u_2 = 0$. This is a contradiction, because the set $\{u_1, u_2\}$ was by assumption a linearly independent set and hence cannot contain a null vector. Hence, we conclude that $u_1 \otimes v_1 + u_2 \otimes v_2 \neq u \otimes v$ for any $u \otimes v \in X \otimes Y$.

Appendix-D

Proposition -11 The range of a separable state $\rho_s : H_1 \otimes H_2 \rightarrow H_1 \otimes H_2$, which is a subspace of $H_1 \otimes H_2$, is spanned by elementary tensors. That is, the subspace $\text{range}(\rho_s)$ has a basis that consists entirely of elementary tensors. For example, if $\rho_s = p_1 P_{x_1 \otimes y_1} + p_2 P_{x_2 \otimes y_2}$, then $\text{range}(\rho_s) = \text{span}\{x_1 \otimes y_1, x_2 \otimes y_2\}$.

Proof: Recall, that the range of a linear operator $T : H \rightarrow H$, is the set $\{T(x) : x \in H\}$; which is a subspace of H . Note, $P_{x \otimes y}(t) = \langle x \otimes y, t \rangle (x \otimes y)$, $t \in H_1 \otimes H_2$.

Let $\rho_s = p_1 P_{x_1 \otimes y_1} + p_2 P_{x_2 \otimes y_2}$, where $p_1, p_2 \geq 0$ and $p_1 + p_2 = 1$. We split the proof into two parts.

Case(i): Assume the set $\{x_1 \otimes y_1, x_2 \otimes y_2\}$ to be linearly dependent. Then, $x_1 \otimes y_1 = \alpha(x \otimes y)$ and $x_2 \otimes y_2 = \beta(x \otimes y)$ for some $\alpha, \beta \in \mathbb{C}$, $x \otimes y \in H_1 \otimes H_2$. Thus, $\rho_s(t) = p_1 \langle x_1 \otimes y_1, t \rangle (x_1 \otimes y_1) + p_2 \langle x_2 \otimes y_2, t \rangle (x_2 \otimes y_2) = p_1 \bar{\alpha} \langle x \otimes y, t \rangle (x \otimes y) + p_2 \bar{\beta} \langle x \otimes y, t \rangle (x \otimes y) = \gamma(x \otimes y)$, $\gamma \in \mathbb{C}$ and $t \in H_1 \otimes H_2$. Hence, the range of ρ_s is the span of the particular element $x \otimes y$, an elementary tensor.

Case (ii): Assume the set $\{x_1 \otimes y_1, x_2 \otimes y_2\}$ to be linearly independent. Now, we claim that there is a $u_0 \in H_1 \otimes H_2$, such that $P_{x_1 \otimes y_1}(u_0) = 0$ and $P_{x_2 \otimes y_2}(u_0) \neq 0$. In that case, $\rho_s(u_0) = p_2 P_{x_2 \otimes y_2}(u_0) = p_2 \langle x_2 \otimes y_2, u_0 \rangle (x_2 \otimes y_2) = \alpha(x_2 \otimes y_2) \neq 0$, and hence $(x_2 \otimes y_2)$ is in the range of ρ_s . Suppose the contrary, that is, assume that there is no such u_0 . This would mean, for any u , for which $P_{x_1 \otimes y_1}(u) = 0$ it follows $P_{x_2 \otimes y_2}(u) = 0$, as well. Recall, that the operator $P_{x \otimes y}$, takes every vector that is orthogonal to $x \otimes y$ to null vector. We denote the set of all vectors that are orthogonal to $x \otimes y$ by $(x \otimes y)^\perp$. Note, if $\dim(H_1 \otimes H_2) = n$, then $(x \otimes y)^\perp = \{t \in H_1 \otimes H_2 : \langle (x \otimes y), t \rangle = 0\}$ is a $(n - 1)$ dimensional subspace of $H_1 \otimes H_2$. Since, $P_{x_1 \otimes y_1}(u) = 0$ implies $P_{x_2 \otimes y_2}(u) = 0$, we have $(x_1 \otimes y_1)^\perp \subset (x_2 \otimes y_2)^\perp$. Observe, $\dim(x_1 \otimes y_1)^\perp = n - 1 = \dim(x_2 \otimes y_2)^\perp$, which implies $(x_1 \otimes y_1)^\perp = (x_2 \otimes y_2)^\perp$. Note, since $\{(x_1 \otimes y_1)^\perp\}^\perp = \text{span}\{x_1 \otimes y_1\}$, and $\{(x_1 \otimes y_1)^\perp\}^\perp = ((x_2 \otimes y_2)^\perp)^\perp$ one concludes that $\text{span}\{x_1 \otimes y_1\} = \text{span}\{x_2 \otimes y_2\}$. This implies that $x_1 \otimes y_1$ and $x_2 \otimes y_2$ are linearly dependent. This is a contradiction. Thus, our claim, that there is a $u_0 \in H_1 \otimes H_2$, such that $P_{x_1 \otimes y_1}(u_0) = 0$ and $P_{x_2 \otimes y_2}(u_0) \neq 0$ is true and hence $(x_2 \otimes y_2)$ is in the range of ρ_s . Reversing the role of $x_1 \otimes y_1$ with that of $x_2 \otimes y_2$, one concludes that $(x_1 \otimes y_1)$ is also in the range of ρ_s . Thus, it is clear that the range of ρ_s is spanned by the elementary tensors $x_1 \otimes y_1$, and $x_2 \otimes y_2$. This proves the proposition.

Appendix-E

Dirac's notation : Let X be a vector space with an innerproduct denoted as $\langle \cdot, \cdot \rangle$ and X^* its dual as defined in section 4.1. In Dirac's notation, $x \in X$ is written as $|x\rangle$, and is called a ket vector and $\phi \in X^*$ is written as $\langle \phi|$, and is called a bra vector. Similarly, what is written as $\phi(x)$, in our notation, where $\phi \in X^*$ and $x \in X$ is written as $\langle \phi|x\rangle$ in Dirac's notation. Right now, the symbol $\langle \cdot, \cdot \rangle$ that occurs in Dirac's notation $\langle \phi|x\rangle$ cannot be interpreted as an innerproduct. This is because the expression $-\langle \phi, x \rangle$ - does not make sense as $\phi \in X^*$ and $x \in X$, live in distinct vector spaces. However, Reisz representation theorem¹¹ says that every continuous linear functional $\phi \in X^*$ can be represented as $\phi(x) = \langle v_\phi, x \rangle$, $x \in X$, where $v_\phi \in X$ is fixed unique vector associated with ϕ . This correspondence, $\phi \in X^* \rightarrow v_\phi \in X$, is a linear map that establishes a one to one

correspondence between X^* and X . The linearity of this correspondence ensures that if ϕ_1 and ϕ_2 are independently mapped to v_1 and v_2 respectively then $\phi_1 + \phi_2$ gets mapped to $v_1 + v_2$. On the other hand, as we saw in section 4.1, every $v \in X$ gets associated with a linear functional ϕ_v , where $\phi_v(x) = \langle v, x \rangle$, $x \in X$. Thus we have a natural (independent of basis) means of identifying elements of X^* with that of X . In other words, this allows us to treat the $\phi \in X^*$ as if it were $v_\phi \in X$ in the sense; $\langle \phi | x \rangle = \phi(x) = \langle v_\phi, x \rangle$, where the last equality makes use of the Reisz representation theorem.

References

¹ Asher Peres, " Quantum Theory : Concepts and Methods, Kluwer Academic Publishers,Chapter-5, 1993.(cf. Chapter-5, Composite systems, Chapter-6, Bell's theorem)

²P.K.Aravind," Quantum mysteries revisited again," Am. J. Phys.**72** (10)1303-1306(2004)

³ R.F.Werner," Quantum information theory- An invitation", In G.Alber, T.Beth, M.Horodecki, R.Horodecki, M.Rotteler, H.Weinfurter,R.F.Werner and A.Zeilinger, " Quantum information: An introduction to basic theoretical concepts and experiments(Springer Tracts in Modern Physics, **173**; Springer-verlag,2001).

⁴ Erling Stormer," Extension of positive maps into $B(H)$ ", Journal of functional analysis, **66** , 235-254 (1986). Lemma-2.2, (page-237) is a generalised version of our Proposition-6. Operator theorists, define a state as a positive linear functionals from the space of operators to complex numbers, that takes the identity operator to the complex number 1. Here, positivity means, the linear functional takes positive operators to positive real numbers. This description is a generalisation of density operators. In the abstract setting of C^* -algebra, classical mechanical states are represented as elements of abelian algebra, while quantum states are from a non-abelian algebra.

⁵ A.Hobson, Concepts in Statistical Mechanics, Gordon and Breach, 1971, Newyork. (cf. Chapter-3, p-53 ; Chapter-4,p-93 to p-100)

⁶ The claim is that it is not possible to represent both position x and momentum p as finite dimensional operators or matrices such that their commutator $xp - px = ihI$. Here I denotes the $n \times n$ identity matrix, $i^2 = -1$ and h is the planck's constant. If it were true, then taking trace on both sides one gets $\text{tr}(xp - px) = \text{tr}(xp) - \text{tr}(px) = 0 = inh$, a contradiction. Recall, $\text{tr}(A + B) = \text{tr} A + \text{tr} B$ and $\text{tr}(AB) = \text{tr}(BA)$.

⁷ when a linear operator T , is represented as a matrix, M_T , then $\text{Tr}(T) = \text{Tr}(M_T) = \text{sum of the diagonal elements of } M_T$. Let $\dim H = 2$

and $B = \{e_1 = (1, 0)^{trp}, e_2 = (0, 1)^{trp}\}$, be the standard orthonormal basis of H , where trp denotes the transpose. Then, $Tr(M_T) = \langle e_1, M_T e_1 \rangle + \langle e_2, M_T e_2 \rangle = (M_T)_{1,1} + (M_T)_{2,2}$, is seen to be the sum of diagonal elements of M_T .

⁸ Let, $T : H \rightarrow H$. Our aim is to show, if y_1 and y_2 are in $\text{range}(T)$, then so is $\alpha y_1 + \beta y_2$ for every $\alpha, \beta \in C$. Since y_1 and y_2 are in the set $\text{range}(T)$, there are vectors x_1 and x_2 in H , such that $T(x_1) = y_1$ and $T(x_2) = y_2$. As T is a linear operator it follows $T(\alpha x_1 + \beta x_2) = \alpha T(x_1) + \beta T(x_2) = \alpha y_1 + \beta y_2$. Thus, if y_1 and y_2 are in the range of T , then so is every linear combination of y_1 and y_2 . Hence, $\text{range}(T)$ is a subspace of H .

⁹ Raymond A. Ryan, "Introduction to tensor products of Banach spaces, " Springer-verlag, 2002. (Chapter-1).

¹⁰ R.F.Werner, " Quantum states with Eienstein-Podolsky-Rosen correlations admitting a hidden variable model," Physical Review A. **40**, 8, 4277-4281 (1989).

¹¹ Martin Schecter, " Principles of functional analysis, ", Academic press.(1971)